AgRISTARS

SR-L0-00425
JSC-16341

A Joint Program for Agriculture and Resources Inventory Surveys Through Aerospace Remote Sensing

May 1980

Supporting Research

TECHNICAL REPORT

A LABELING TECHNOLOGY FOR LANDSAT IMAGERY

T. B. Dennis and M. D. Pore

LOCKHEED ENGINEERING AND MANAGEMENT SERVICES COMPANY, INC.
1830 NASA Road 1, Houston, Texas 77058

SR-LO-00425
JSC-16341

TECHNICAL REPORT

A LABELING TECHNOLOGY FOR LANDSAT IMAGERY

Job Order 73-302

This report describes classification activities of the Supporting Research project of the AgRISTARS program.

PREPARED BY

T. B. Dennis and M. D. Pore

APPROVED BY

J. C. Minter

T. C. Minter, Supervisor
Techniques Development Section

J. E. Wainwright

J. E. Wainwright, Manager
Development and Evaluation Department

LOCKHEED ENGINEERING AND MANAGEMENT SERVICES COMPANY, INC.
Under Contract NAS 9-15800

For

Earth Observations Division

Space and Life Sciences Directorate

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
LYNDON B. JOHNSON SPACE CENTER
HOUSTON, TEXAS

May 1980

LEMSCO-14357

CONTENTS

TABLES

FIGURES

1. INTRODUCTION

1.1 FIELD DELINEATION AND LABELING

In the Large Area Crop Inventory Experiment (LACIE), Landsat imagery was analyzed in an effort to monitor the world-wide production of wheat. To estimate the wheat production in a given region, several 8- by 9-kilometer (5- by 6-mile) segments located within the region were extracted from the Landsat data. Individual acreage estimates were made for each segment. These acreage estimates were then aggregated to obtain Crop Reporting District (CRD) acreage estimates which, in turn, were multiplied by CRD yield estimates to obtain production estimates. A large source of variance in this procedure lies in the acreage estimation of the individual segments.

In LACIE Phases I and II (1975 and 1976 growing seasons), acreage estimates were made by performing a maximum likelihood classification of the picture elements (pixels) in each segment. This process assumes that the data follow a mixture of Gaussian distributions. Samples are required in estimating the particular mixture present in the scene. The individual pixels are then classified as belonging to the most likely distribution, based on the pixel's spectral values and the mixture distribution estimated from the observed samples. Throughout Phases I and II of the experiment, analyst interpreters (AI's) gathered and labeled the samples necessary for this procedure.

To obtain the necessary samples, the AI observed film products that were generated from the Landsat data. The AI's job was to choose and label representative samples from the scene which, it was assumed, constituted a mixture of normal distributions. In choosing the samples, the AI observed the imagery and selected and delineated fields within the image. The task involved the sampling of all major underlying distributions in proportion to their representation in the scene. Once the samples were chosen, the AI used the imagery in conjunction with ancillary data to provide corresponding labels.

The training and classification described above was normally done using a single 4-channel acquisition of a Landsat segment. Some segments were processed multitemporally, but it should be noted that the problem of sampling all major distributions in the correct proportions greatly increased with added acquisitions. Therefore, to benefit from the added acquisitions used for identifying confusion crops, the AI had to accept the drawback of compounding the training problem and increasing the time required for processing.

In addition to the AI problems of choosing the acquisition or acquisitions to process and choosing a representative training sample, the field delineation approach had other drawbacks. For example, the sample of each underlying distribution was generally inadequate in that the extremes of the distribution were rarely sampled. Also, in areas where crops were grown in small fields, there was often difficulty in obtaining a reliable sample of each signature. Another problem noted with this approach was in its inefficient use of AI resources. Of the total time spent by the analysts in processing, approximately one-eighth of that amount was spent in performing the most important task, the labeling of the samples. To overcome these difficulties, a procedure based on the sampling and labeling of individual pixels, known as Procedure 1, was developed at the beginning of Phase III (the 1977 growing season).

1.2 PIXEL LABELING: ACQUISITION USAGE

As a replacement for field delineation, a clustering algorithm was employed in Procedure 1 to produce training samples. In this procedure, the AI was required to label a random sample of pixels from each segment. A subset of this sample, called type 1 dots, was used to seed the clustering algorithm. Only those type 1 pixels which sampled the same field on all acquisitions were used. The associated labels were used to label the output clusters according to a nearest neighbor rule. These labeled clusters were then used as training samples for the maximum likelihood classifier. The remaining pixels of the original random sample, called type 2 dots, were used to compute a stratified random proportion estimate from the strata produced by the classifier. Type 2

dots were not required to sample the same fields on each acquisition; but they were labeled on the basis of their location on a specified base acquisition.

The use of a random sample of the scene to produce clusters was intended to remove any variance that could be caused by biases in the field delineation method of sample selection. Furthermore, this method had the advantage of allowing the use of multiple acquisitions without increasing the work required to extract a representative training sample. The role of the analyst was thus reduced to that of selecting a set of up to four acquisitions which best characterized the separation between small grains (wheat, barley, oats, and flax) and nonsmall grains and labeling the random sample of pixels represented in those acquisitions.

At this stage in the procedure, a labeling technology was needed to reduce the variance associated with the AI labeling.

1.3 AI QUESTIONS DEVELOPED

In order to analyze the relative importance of the various factors comprising an AI interpretation, a list of questions relating to these factors was compiled by a team of experienced AI's. The questions related to agricultural practices, meteorological conditions, and spectral values that influence pixel analysis, as well as subjective film product interpretation regarding the field membership and vegetation canopy of certain pixels. The questionnaire described the interpretation of pixel labels used in LACIE. The required responses to some of the questions were qualitative: yes, no; bad, good; better, best; etc. Other questions required quantitative answers: amount of rainfall in inches, various transformations of the radiometric spectral values, etc. The qualitative responses were coded with nonnegative integer values, and a vector of all responses was composed for each pixel.

Four 8- by 11-kilometer (5- by 6-mile) segments were analyzed using a grid of 209 pixels. The grid consisted of every tenth pixel, both horizontally and vertically. This 10-by-10-grid was the same grid introduced with Procedure 1 to eliminate pixel-to-pixel (interfield) dependencies in spectral values and

interpretation. To develop a more objective procedure, the AI opinions regarding any small-grain-versus-"other" labeling were ignored in the analysis of the questions, and a discriminate analysis was applied to the vector of the AI responses to differentiate pixels that were members of the ground-truth small-grain category from the members of the "other" category. The intention was to imitate the procedure the AI followed in weighting various sources of information to determine pixel labels. It was also desired that the procedure would provide an estimate of the accuracy of these labels. Due to a shortage of data, the classifications produced by the discriminant analysis were tested on the training data rather than on a separate test set. Using the results of these tests, repeated discriminate analyses were generated step by step; and, in conjunction with AI consultations concerning the logic of the interpretation process, a succinct set of key questions that would not significantly sacrifice classification accuracy was generated. This set of key questions, along with the procedure for its use, has been named Label Identification from Statistical Tabulation (LIST). The LIST questions were partitioned into two groups: spectral questions (for which responses were computed directly by the computer) and AI questions (for which answers were obtained from analyst interpretations). The automation of the spectral information was important in producing an operationally feasible pixel-labeling procedure that is cost effective in terms of interpretation time.

The LIST questions and analysis procedure used in the experiment are described in the following section. Experimental results (both training and test) concerning the accuracy of labeling are discussed in section 3.

2. LIST

List data consist of two parts, the part acquired from the AI and the automated part derived from the spectral values. In accordance with the LIST procedure, the AI is given a packet that contains all available film products, agricultural-meteorological background data, and appropriate maps for a given area. From the available film imagery, the AI selects four available acquisition dates for the interpretation. The chosen dates are selected because they span the growing season of the crop of interest (spring wheat in North Dakota, for example) and reflect key stages of growth, such as heading (peak vegetation canopy) and harvest (no vegetation). Each acquisition is assigned an average biostage rating using the Robertson biostage scale (ref. 1), which is adjusted for local weather conditions during the growing season. All crops of interest in the scene are expected to be within one biostage of the average biostage rating assigned for that particular acquisition.

The AI interprets the pixels on the film imagery to provide AI pixel-specific responses to the questions in the questionnaire shown in figure 2-1. These responses are recorded on an AI response sheet (see figure 2-2) in a format suitable for keypunching. Notice that the segment identification number, the acquisition dates, and the respective Robertson biostage numbers are recorded on the top line of the AI response sheet. The sixth response in the questionnaire (figure 2-1), the AI interpretation, calls for an answer based on the AI's training, experience, visual acuteness, and the amount of time and care taken by the AI in making a study of the vegetation patterns in the segment. The variety of responses given by the analysts indicates that, in many cases, the evaluations made are highly subjective. This response is not used in the first part of the LIST procedure, but it is used to identify possible problem pixels later in the procedure.

The responses indicated on figure 2-2 are key punched and represent one data source for the LIST computer software. The other data source is a tape of the Landsat multispectral scanner (MSS) radiometric values for each pixel in the

AI PIXEL-SPECIFIC RESPONSES

1-4 FOR EACH ACQUISITION.
PFC VEGETATION CANOPY INDICATION IS _____.
(USE ALL AVAILABLE IMAGERY FILM TYPES.)

(0) NO VEGETATION CANOPY
(1) LOW DENSITY GREEN VEGETATION CANOPY
(2) MEDIUM DENSITY GREEN VEGETATION CANOPY
(3) HIGH DENSITY VEGETATION CANOPY
(4) SENESCING (TURNING) VEGETATION CANOPY
(5) HARVESTED CANOPY (STUBBLE)

5 THE MULTITEMPORAL ORIENTATION OF THE PIXEL ACROSS THE FOUR ACQUISITIONS IS _____.

(D) DESIGNATED OTHER: OBVIOUSLY IN A NONAGRICULTURAL AREA, NOT IN A FIELD
(R) REGISTRATION: PIXEL SWITCHES FIELDS
(M) MIXED: PIXEL IS NOT ENTIRELY IN ONE FIELD
(P) PURE: PIXEL IS IN THE SAME FIELD ON ALL FOUR ACQUISITIONS

6 THE AI INTERPRETATION OF THE PIXEL CATEGORY IS _____.

Figure 2-1.- Sample AI questionnaire form used in the analysis procedure.

scene. This latter data source is screened to admit only those pixels interpreted by the AI. The MSS data set for each pixel is a 16-dimensional vector representing light reflectance in the green, red, near infrared, and far infrared bands, respectively, for each of the four acquisitions.

The LIST program first transforms the AI responses and MSS data into variables that relate to the growth stages for the crop in question. The program then transforms those responses and data to weight each variable according to its contribution in the decision making process as determined by the training data. The scalar sum of the weighted responses then reflects the degree of confidence one can place on the classification. For this process of discriminant analysis, training samples are required in order to determine the weighting and threshold for classifications using the weighted sum. In the data analysis presented in the following section, the training of the discriminator is discussed and illustrated. First, however, an explanation of the transformation of analyst responses and MSS data into the LIST keys is given.

The AI vegetation canopy responses shown in figure 2-2 are used in conjunction with the data in figure 2-3 to determine a variable called the "canopy key." As shown in figure 2-3, each acquisition's biostage is noted on the horizontal axis, and the vegetation canopy code is noted on the vertical axis. This figure has been generated to accommodate the growing phase of wheat and other small grains in the U.S. Great Plains. If a pixel is plotted into the blank area (in the middle), it is considered a "first class" response for small grains and its canopy key is 0. If it is plotted into the dotted region (next to the blank area), it is considered a marginal response and its canopy key is 5. If it is plotted into one of the slashed regions (upper left or lower right), it is considered an unacceptable response for small grains and its canopy key is 10. Canopy keys are determined for each acquisition of each pixel. An additional variable, called the "canopy trajectory," is generated by summing the canopy keys and setting the canopy trajectory equal to 0 if the sum is less than or equal to 5 and equal to 1 if the sum is greater than or equal to 10. The canopy keys will be denoted CANKY(I,J), J = 1, 4, where I is

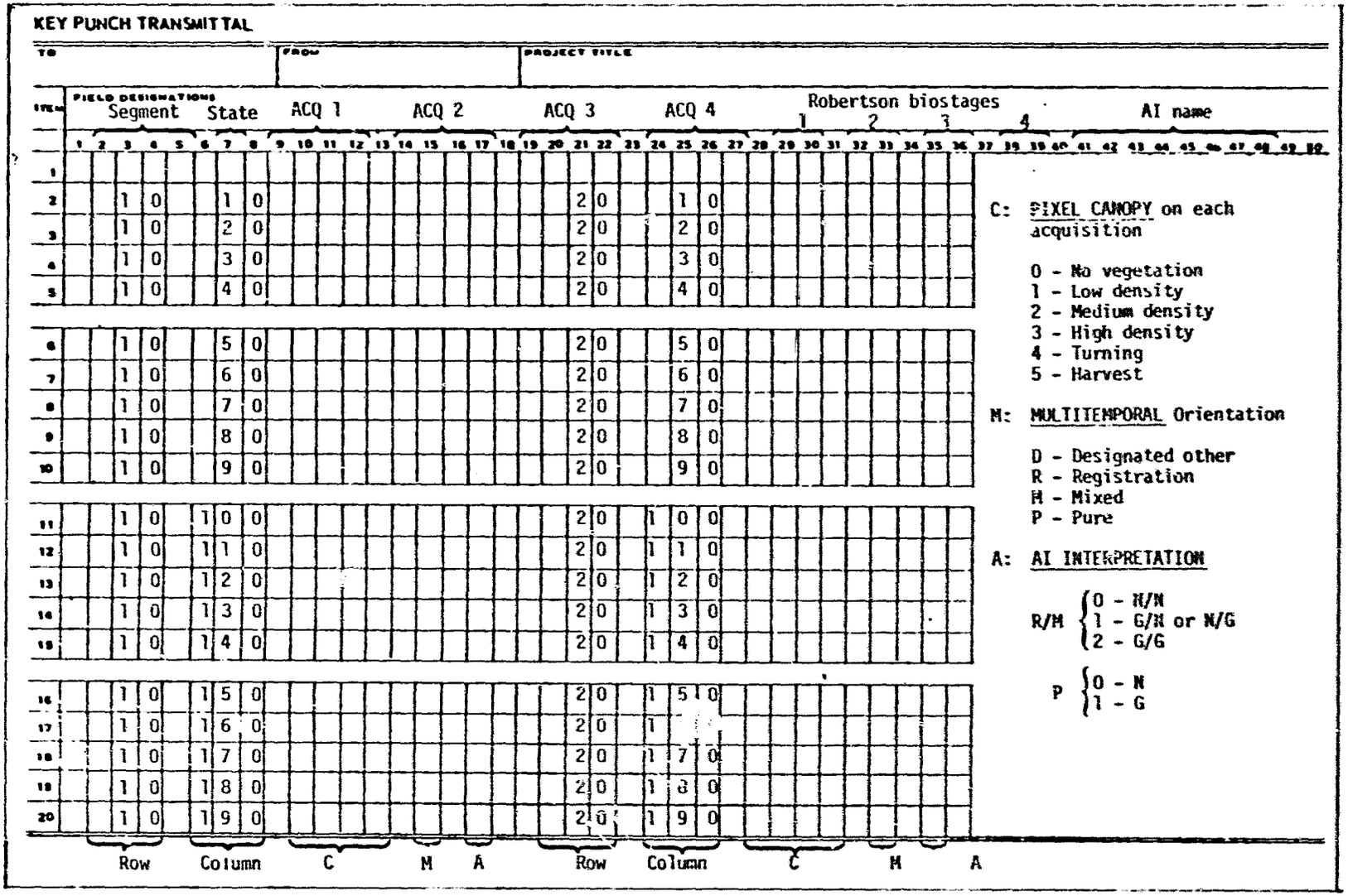


Figure 2-2.- Sample AI response sheet.

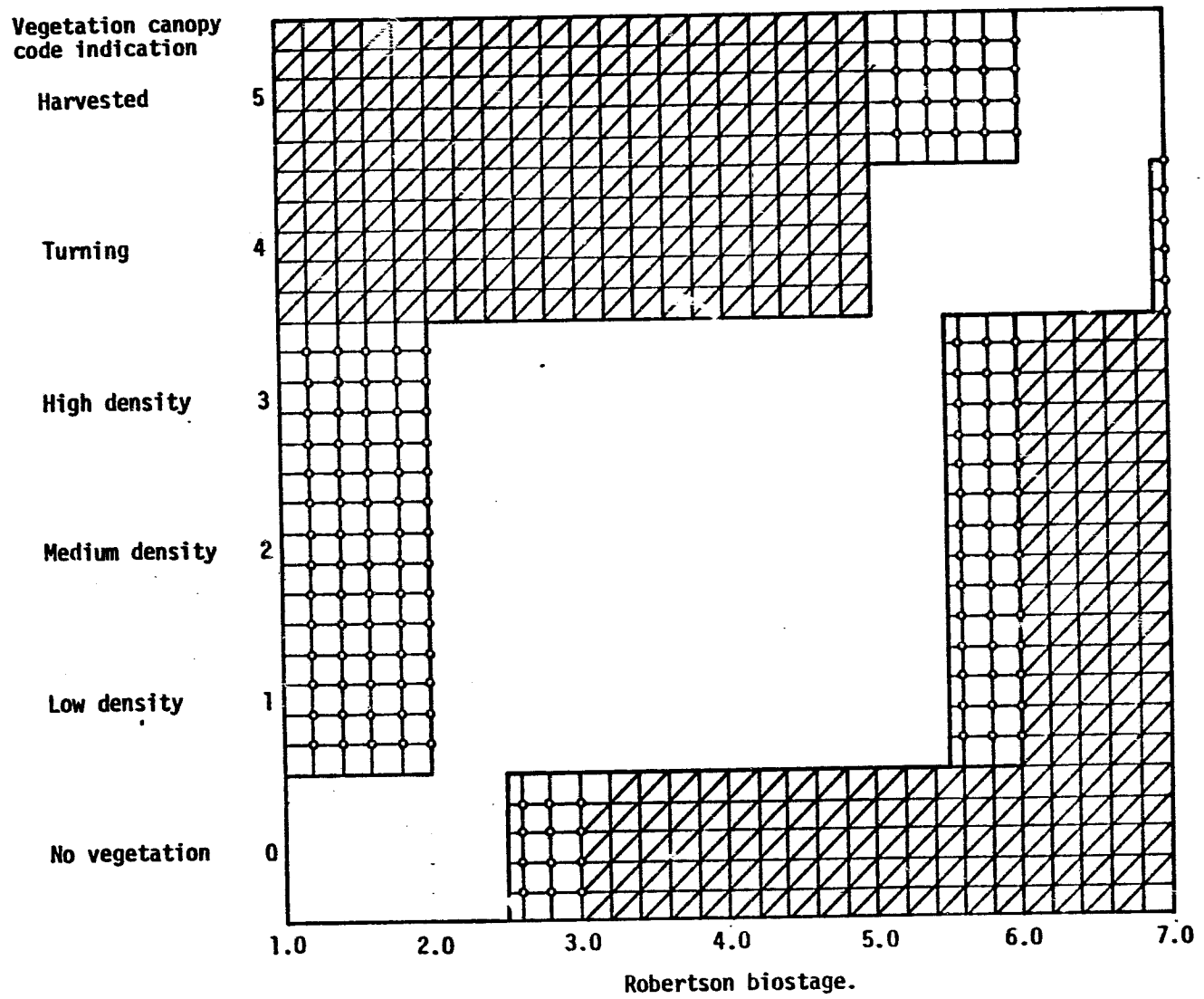


Figure 2-3.- Illustration showing AI vegetation canopy responses.

an index over the interpreted pixels, J is an index to the acquisition number, and the canopy trajectories are denoted CANTJ(I).

The recoding of the spectral values is a little more complex. All of the spectral variables are transformations of greenness and brightness. Greenness and brightness are, in turn, linear transformations of the 4-dimensional MSS radiometric values. [See Kauth and Thomas (ref. 2) for a physical interpretation of greenness and brightness.] For the crop of interest, a prototype or "expected trajectory" in each of the greenness (GREEN) and brightness (BRIET) dimensions is generated along with an empirical standard deviation of the estimator. Specific generation techniques used may vary according to local conditions. In section 4, these techniques, as well as that used for the test described in section 4, are explained.

The biostage means and standard deviations are used to form "z-scores" (observed scores) for each pixel on each acquisition, as follows:

$$\text{BRIET}(i,j) = [B(i,j) - \text{MEANB}]/\text{SDB}$$

where

i = pixel (1-209)

j = index to acquisition number

B(i,j) = brightness value extracted from 4-dimensional vector of acquisition j

MEANB = mean of brightness

SDB = standard deviation of brightness

and

$$\text{GREEN}(i,j) = [G(i,j) - \text{MEANB}]/\text{SDG}$$

where

i = pixel (1-209)

j = index to acquisition number

G(i,j) = greenness value extracted from 4-dimensional vector of acquisition j

MEANB = mean of brightness

SDB = standard deviation of brightness

The variables denoted BRIET(i,j) and GREEN(i,j) are concatenated with CANKY(i,j) and CANTJ(i) to form a larger vector. This vector is then augmented with the absolute z-scores and four additional trajectory variables as follows:

$$ABREIT\ (I,J) = |BRIET(I,J)|$$

$$AGREEN(I,J) = |GREEN(I,J)|$$

$$SQAIRB(i) = \sum_{j=1}^{4} [BRIET(i,j)^2]$$

$$SQUAIRG = \sum_{j=1}^{4} [GREEN(I,J)^2]$$

$$PIEB(i) = \prod_{j=1}^{4} [1 + ABRIET(i,j)]$$

$$PIEG(i) = \prod_{j=1}^{4} [1 + AGREEN(i,j)]$$

where ABRIET is the absolute value of brightness, and AGREEN is the absolute value of greenness. The vector of LIST keys is now a 25-dimensional vector. This is the vector on which the discriminant analysis is based.

The weightings for each variable can be derived in various ways. In this study, weights were derived by using a classical discriminant procedure in which, for the segments of interest, known (ground-truth) labels were observed. Let us assume that, for the particular area to be interpreted, an appropriate set of weights has been determined, perhaps through the use of discriminant coefficients trained on the previous year's data. The 25-dimensional supervector is then converted to a single discriminant score by applying the weights and summing. Zero is the natural threshold for classification when discriminant coefficients are used. The result is a classification for each pixel that the interpreter analyzed. These discriminator labels

TYPE 1 DOT CLASSIFICATION

	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150	160	170	180	190
10		S/S		S/S		S/S		R/S		S/S		S/S		S/S		D/N		M/N	
20	S/S		S/S		D/N		M/S		M/S		D/N		S/S		S/S		D/N		N/N
30		M/S		S/S		D/N		D/N		D/N		S/S		D/N		D/N		D/N	
40	D/N		M/N		S/S		S/S		R/S		D/N		R/N		D/N		D/N		R/N
50		N/S		N/S		D/N		M/S		M/S		N/S		D/N		M/S		N/N	
60	N/S		N/S		M/S		M/S		S/S		S/S		M/S		D/N		N/S		S/S
70		D/N		S/S		N/S		N/S		N/S		S/S		M/N		R/N		D/N	
80	S/S		D/N		R/N		M/S		N/N		S/S		D/N		M/N		N/N		D/N
90		D/N		N/S		N/S		M/S		N/S		N/S		D/N		R/S		M/N	
100	D/N		S/S		M/N		M/S		N/N		N/S		N/N		N/N		N/N		N/N
110		N/N		S/S		S/S		S/S		N/N		M/S		N/N		R/N		M/S	

11 2-8

Figure 2-4.- Analyst-interpreted labels versus discriminator labels, type 1 dots.

are then arrayed along with the AI opinion given in the last question in the LIST analysis (see figure 2-4 for array). The interpreter examines those pixels over which disagreements have occurred. The procedure used in this analysis was to consider the discriminator labels as final, unless the interpreter could state a reason for preferring his label. Making a change in the discriminator label is acceptable when, for example, additional acquisitions show growth of a crop which was not evident in the four acquisitions used or the previous year's data indicate agricultural practices which predict growth of a particular crop for the current year.

Thus, the LIST labeling procedure is a technology that uses the interaction of both the automated discrimination techniques and the photointerpretation experience in deriving pixel labels. It enables the interpreter to work without the continual use of confusing or difficult spectral aids. The numerical results of the use of LIST on data collected from N. Dakota blind sites in LACIE Phase III (1977) and the 1978 Transition Year (TY) growing season are given in the following section.

3. EXPERIMENTAL RESULTS: N. DAKOTA

3.1 TRAINING RESULTS IN THE 1976-77 DATA

To show that the LIST procedure can be made operational, an experiment was devised. LIST was trained on Phase III (1976-77 growing season) spring small-grain data from N. Dakota to obtain a discriminant function. This discriminant was applied to the N. Dakota spring small-grain data collected in Phase III to estimate the training accuracy of the procedure and to N. Dakota, S. Dakota, and Minnesota data collected in the TY (1977-78 growing season) to estimate the temporal and geographic extendability of the procedure.

The first step in training LIST for use in a specific geographic area is to obtain the expected greenness and brightness trajectories of small grains used to transform the MSS data to LIST spectral keys. In this experiment, the trajectories were obtained from the available ground-truth small-grain pixels in the N. Dakota Phase III data. The pixels were taken from the 14 blind sites which had the necessary four acquisitions required by LIST, though, in general, this is not a necessary restriction for generating the trajectories. The pixels were treated as four independent observations, with one observation on each acquisition. The acquisitions were first divided into groups, with each group consisting of all the acquisitions obtained during one 18-day cycle of Landsat coverage. The range of the Robertson biostage occurring within each group was noted. The means and standard deviations of greenness and brightness were computed for each group. The expected trajectories of greenness and brightness as a function of the Robertson biostage were then generated by applying the observed means to the appropriate biostages and linearly interpolating to cover unobserved biostages. This procedure was repeated to determine the standard deviation of greenness and brightness for each biostage. The resulting trajectories are presented in figures 3-1 and 3-2.

With these trajectories computed, the AI responses and MSS data from the 14 Phase III N. Dakota blind sites were transformed to the 25-dimensional LIST keys. A discriminant was trained to separate the ground-truth small grain pixels from the ground-truth "other" pixels represented by these transformed

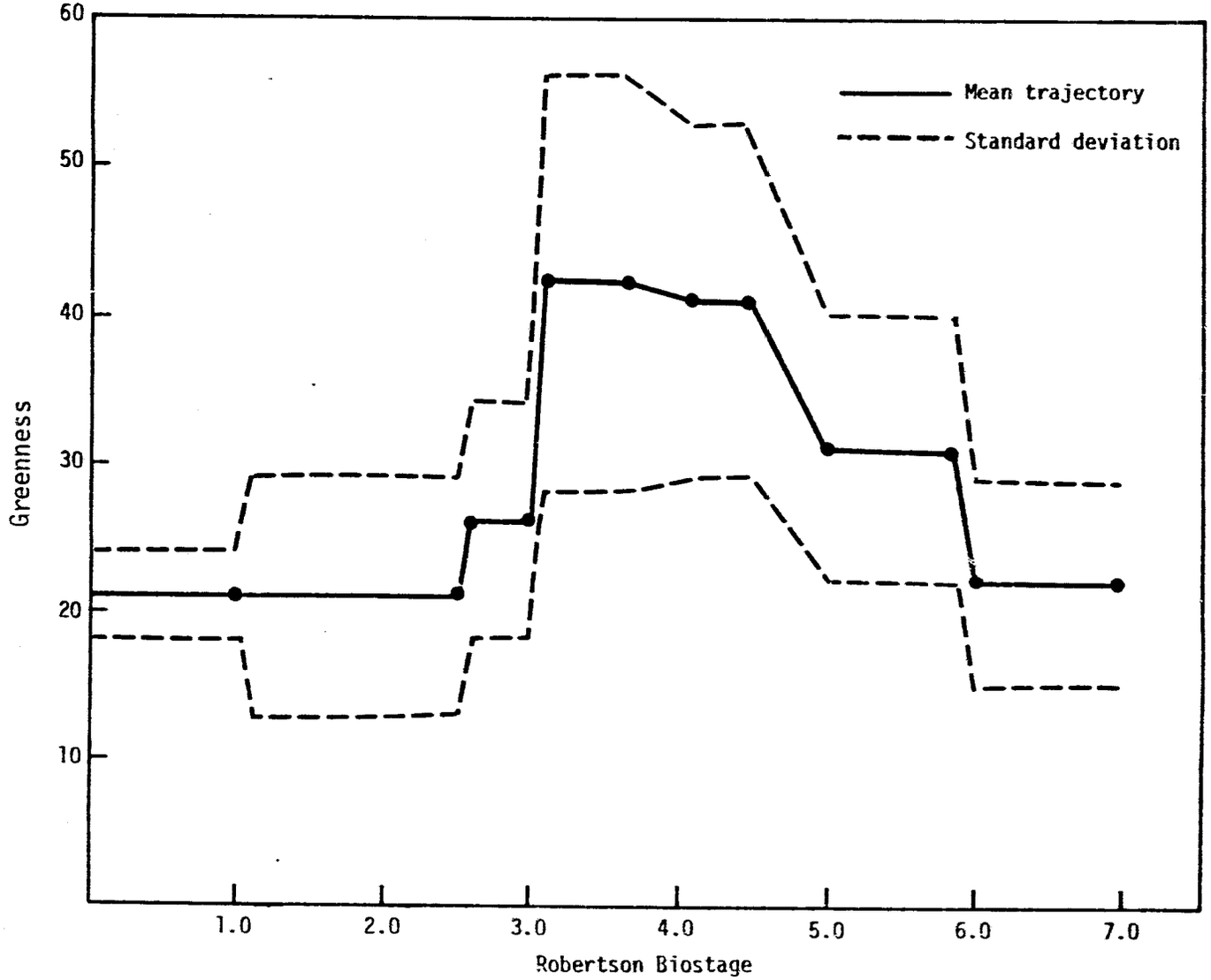


Figure 3-1.- Phase III greenness key generated from the ground-truth labels for North Dakota segments.

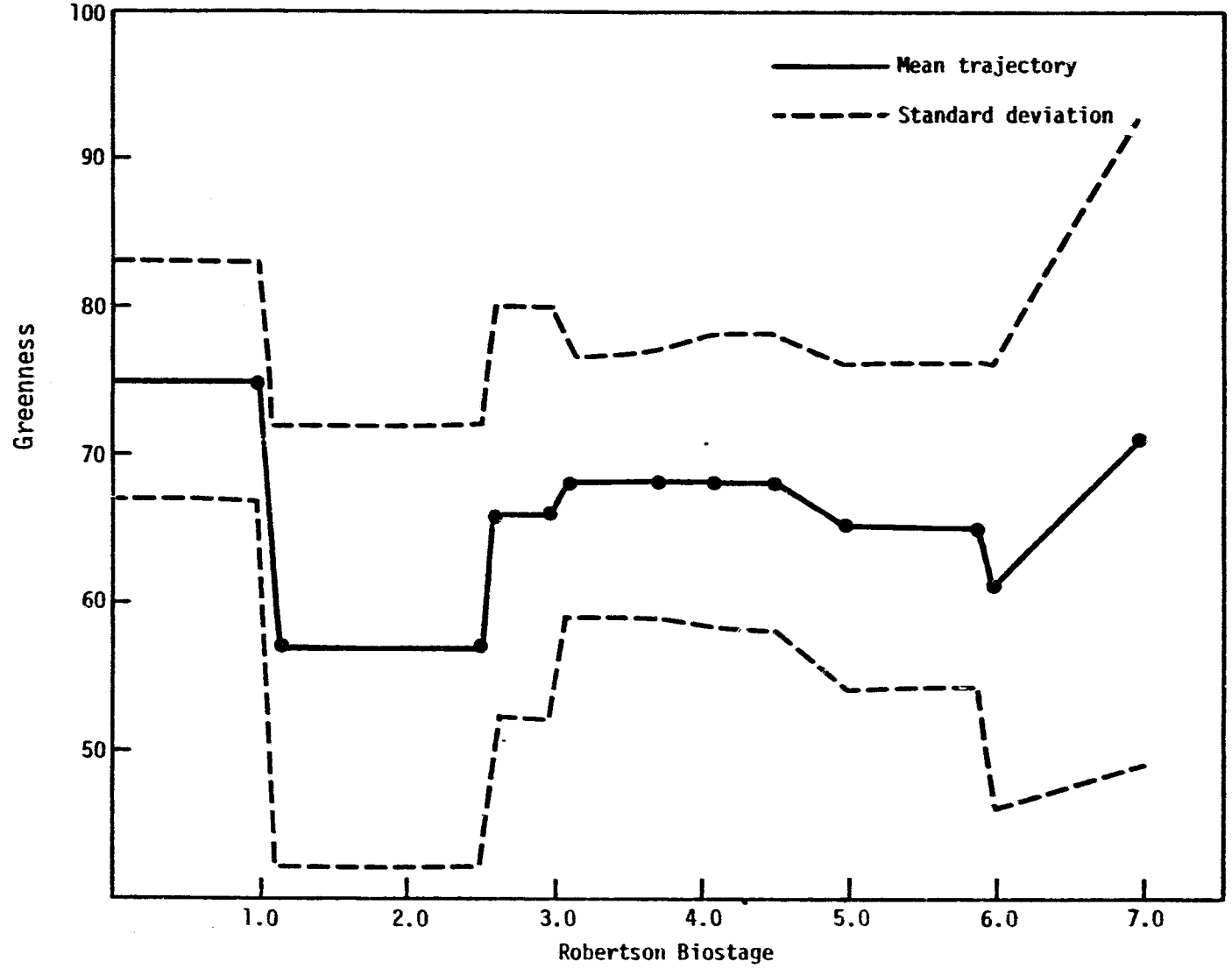


Figure 3-2.- Phase III brightness key generated from the ground-truth labels for North Dakota segments.

response vectors. Table 3-1 shows the labeling accuracy obtained by applying this discriminant to the same data set, and it shows the accuracy of the analyst label (provided as a response to LIST by each analyst) for comparison. In this table, PCC stands for probability of correct classification. It is computed as the number of ground-truth small grain pixels classified as small-grains plus the number of ground-truth "other" pixels, either classified as "other" or labeled "obviously nonagriculture," divided by the total number of pixels. The omission rate is the percentage of ground-truth small grains that were not classified as small grains, and the commission rate is the percentage of ground-truth "other" pixels that were classified as small grains. The remainder of the table is self-explanatory.

3.2 TEST RESULTS

The next step in the test of the LIST procedure was to apply the Phase III N. Dakota discriminant to the following year's data from N. Dakota, S. Dakota, and Minnesota. This provided a twofold test of temporal and geographic extendability of the procedure. The results of this test are shown in tables 3-2 and 3-3. Table 3-2 shows the initial results. In both cases the accuracy was low. The fact that the discriminant did not provide better accuracy in N. Dakota than in the other states indicates that the chief problem was the temporal rather than the geographic extension. Evidence to support this conclusion is given in table 3-3, which shows the results obtained by training on the N. Dakota TY data and geographically extending the discriminant to six additional S. Dakota and Minnesota segments. A study of the causes of this poor temporal extension was made, and an evaluation of the results is included in the next section of this document.

3.3 EVALUATION OF RESULTS

The first attempt to improve the temporal extension of the LIST labeling technology involved temporally updating the spectral keys used in the procedure. It was necessary to achieve this without the benefit of the ground truth in order to maintain an operational procedure for labeling in a situation where the ground-truth data were unavailable. This method used the AI labels that

TABLE 3-1.- TRAINING RESULTS FOR PHASE III NORTH DAKOTA SEGMENTS

(a) Distribution of LIST labels

Ground-truth label	LIST label		
	Small grains	Nonsmall grains	Obvious nonagriculture
Small grains	534	167	13
Nonsmall grains	143	669	496

Statistics:
- PCC = 84.07%
- Omission rate = 25.21%
- Commission rate = 10.93%
- Bias = -1.8%
- Average PCC across segments = 84.31%
- Standard deviation of PCC = 4.69%
- PCC, given LIST and AI agree = 88.03%
- PCC of LIST on disagreements = 40.97%

(b) Distribution of AI labels

Ground-truth label	AI label		
	Small grains	Nonsmall grains	Obvious nonagriculture
Small grains	370	330	13
Nonsmall grains	63	751	496

Statistics:
- PCC = 80.00%
- Omission rate = 48.11%
- Commission rate = 4.66%
- Bias = -14.0%
- Average PCC across segments = 80.46%
- Standard deviation of PCC = 9.75%

TABLE 3-2.- INITIAL RESULTS FROM CLASSIFYING TY DATA WITH THE PHASE III DISCRIMINANT

(a) Distribution of LIST labels for 19 TY sites in North Dakota, South Dakota, and Minnesota

Ground-truth label	LIST label		
	Small grains	Nonsmall grains	Obvious nonagriculture
Small grains	339	612	12
Nonsmall grains	660	1005	246

Statistics:
- PCC = 55.32%
- Omission rate = 64.00%
- Commission rate = 34.54%
- Bias = +1.7%
- Average PCC across segments = 57.13%
- Standard deviation of PCC = 20.14%
- PCC, given LIST and AI agree = 81.07%
- PCC of LIST on disagreements = 18.79%

(b) Distribution of LIST labels for 14 North Dakota TY blind sites

Ground-truth label	LIST label		
	Small grains	Nonsmall grains	Obvious nonagriculture
Small grains	286	512	9
Nonsmall grains	406	797	110

Statistics:
- PCC = 52.26%
- Omission rate = 63.44%
- Commission rate = 30.97%
- Bias = -5.0%

TABLE 3-3.- TRAINING RESULTS FOR TY NORTH DAKOTA, SOUTH DAKOTA, AND MINNESOTA DATA

(a) Distribution of LIST labels for 15 North Dakota blind sites

Ground-truth label	LIST label		
	Small grains	Nonsmall grains	Obvious nonagriculture
Small grains	502	323	10
Nonsmall grains	128	1230	196

Statistics:
- PCC = 80.70%
- Omission rate = 39.80%
- Commission rate = 8.20%
- Bias = -8.5%
- Average PCC across segments = 79.23%
- Standard deviation of PCC = 11.22%
- PCC, given LIST and AI agree = 83.6%
- PCC of LIST on Disagreements = 54.1%

(b) Distribution of LIST labels for 21 North Dakota, South Dakota, and Minnesota blind sites

Ground-truth label	LIST label		
	Small grains	Nonsmall grains	Obvious nonagriculture
Small grains	583	418	14
Nonsmall grains	127	1788	322

Statistics:
- PCC = 82.81%
- Omission rate = 42.56%
- Commission rate = 5.68%
- Bias = -9.4%
- Average PCC across segments = 84.27%
- Standard deviation of PCC = 11.5%

were supplied to the LIST processor as a substitute for ground-truth labels in the generation of the trajectories. The trajectories determined in this way are shown in figures 3-3 and 3-4. They are not significantly different from the corresponding trajectories generated from the ground-truth labels (figures 3-5 and 3-6) and, therefore, this method of updating trajectories was adopted. Table 3-4 shows the results obtained by substituting the updated trajectories in the processor that generated the LIST keys. Since the improvement obtained by this process was minimal, a further study was made of the contribution of individual keys to the problem.

The two sets of keys which contributed the most to the lack of temporal extendability were found to be the brightness keys and the analyst keys. Table 3-5 shows the test results obtained by (1) removing the brightness keys, (2) training in Phase III, and (3) classifying the TY data. The increase in accuracy and the significant changes in the brightness trajectory (figures 3-2, 3-3, and 3-4) from Phase III to the TY indicate that the brightness keys are unstable. Tables 3-6 and 3-7 show the mean PCC's and standard deviations for training segments in Phase III with test segments in the TY. Table 3-6 shows the results obtained by using only the greenness and brightness keys. Table 3-7 indicates the results obtained by using only the greenness keys. The improvement obtained by removing the AI keys was considered significant. Table 3-8 shows the results obtained using only the analyst keys. The fact that the Phase III discriminant obtained from these keys explained only 56 percent of the Phase III analyst labeling indicates that a problem existed in the AI responses collected for the Phase III data. It is believed that the problem occurred because only two AI's were available at the time to support the collection of this data. By contrast, the broad set of responses obtained from using 16 AI's to interpret the TY data produced a discriminant which explained 87 percent of the AI labeling (table 3-8). Finally, table 3-9 indicates the mean PCC and standard deviations for training and test data from the TY, showing again the relatively good geographic extension that can be obtained using the LIST procedure.

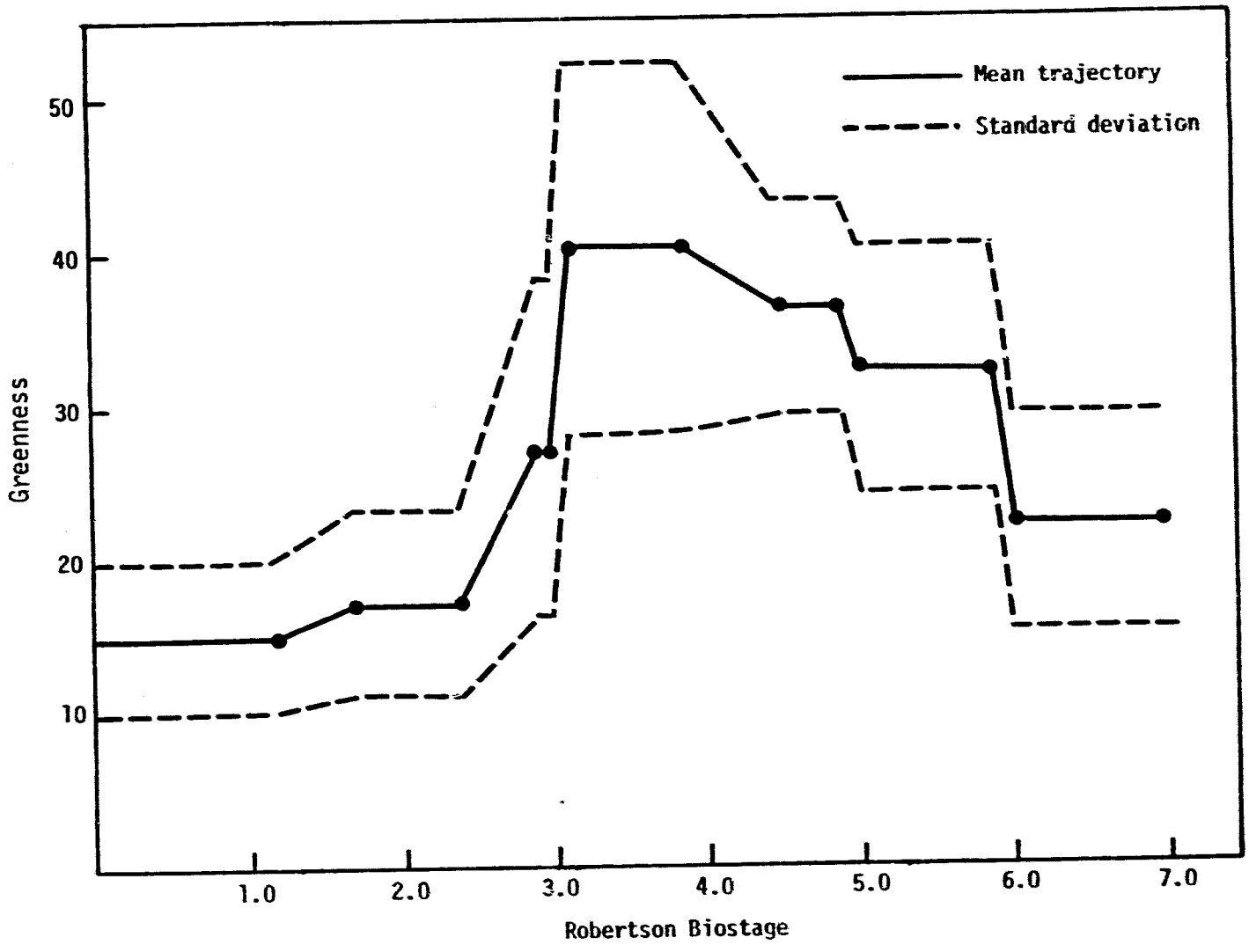
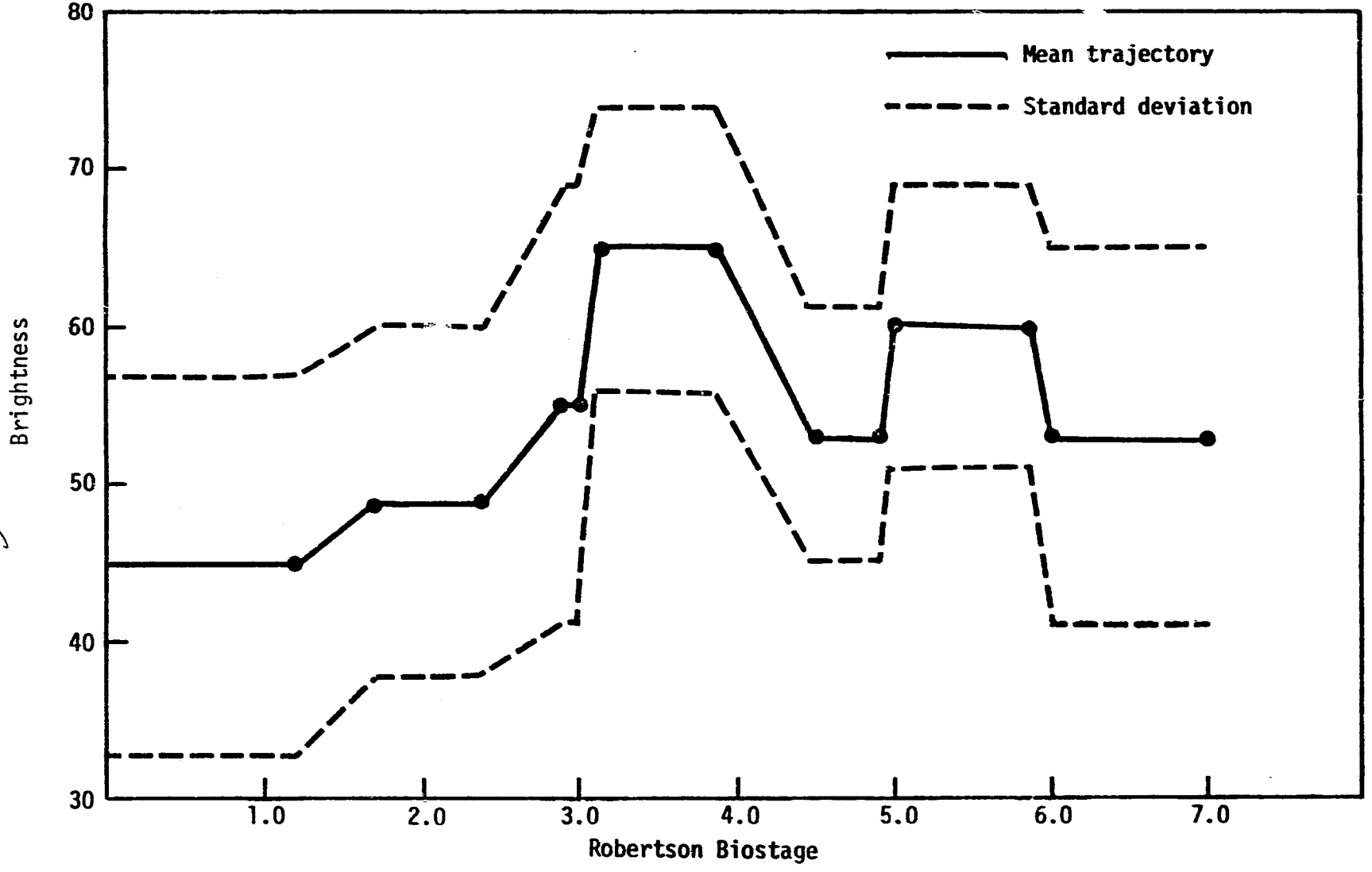


Figure 3-3.- TY greenness key generated from the AI labels.



3-10
23

Figure 3-4.- TY brightness key generated from the AI labels.

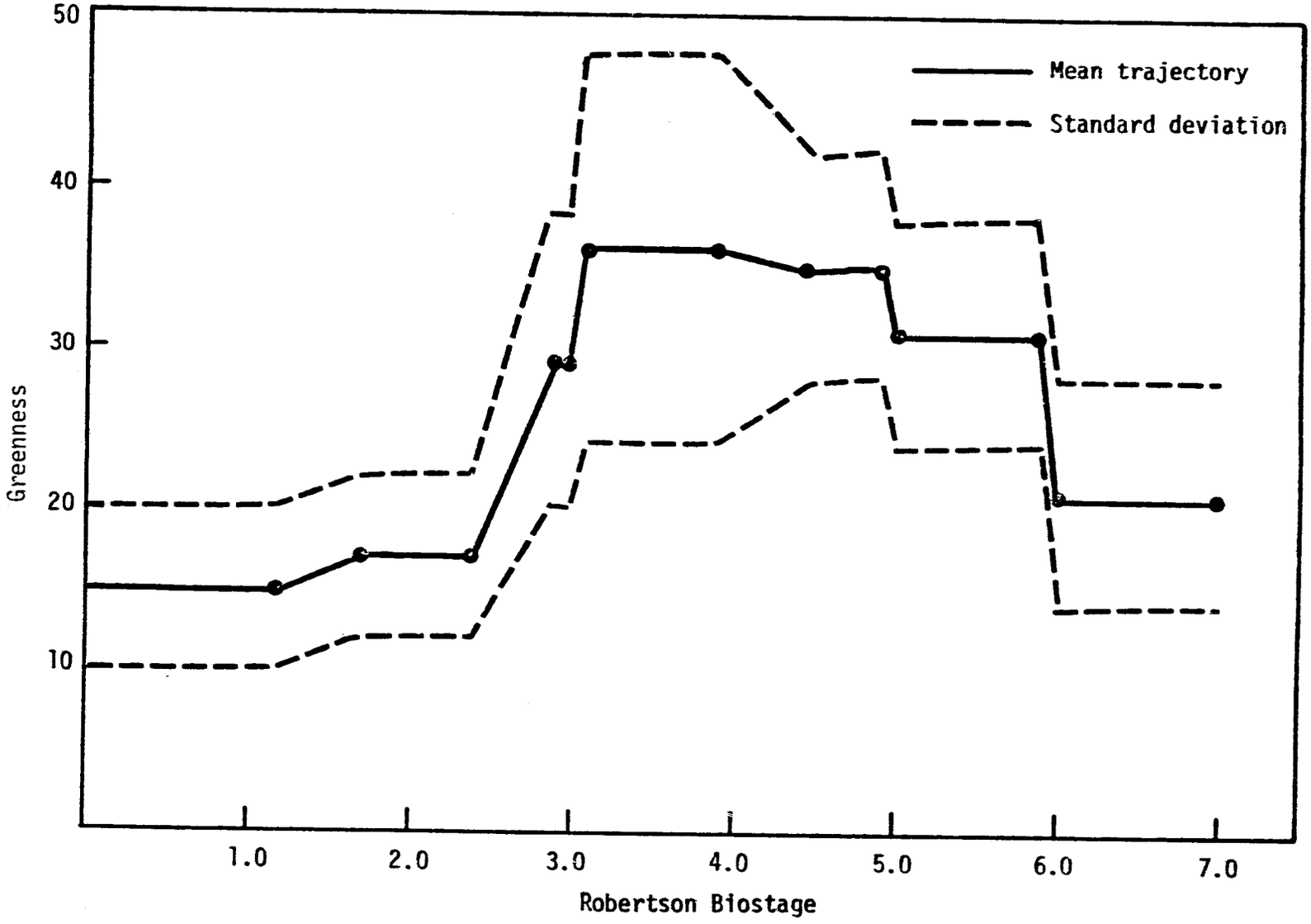


Figure 3-5.- TY greenness key generated from the ground-truth labels.

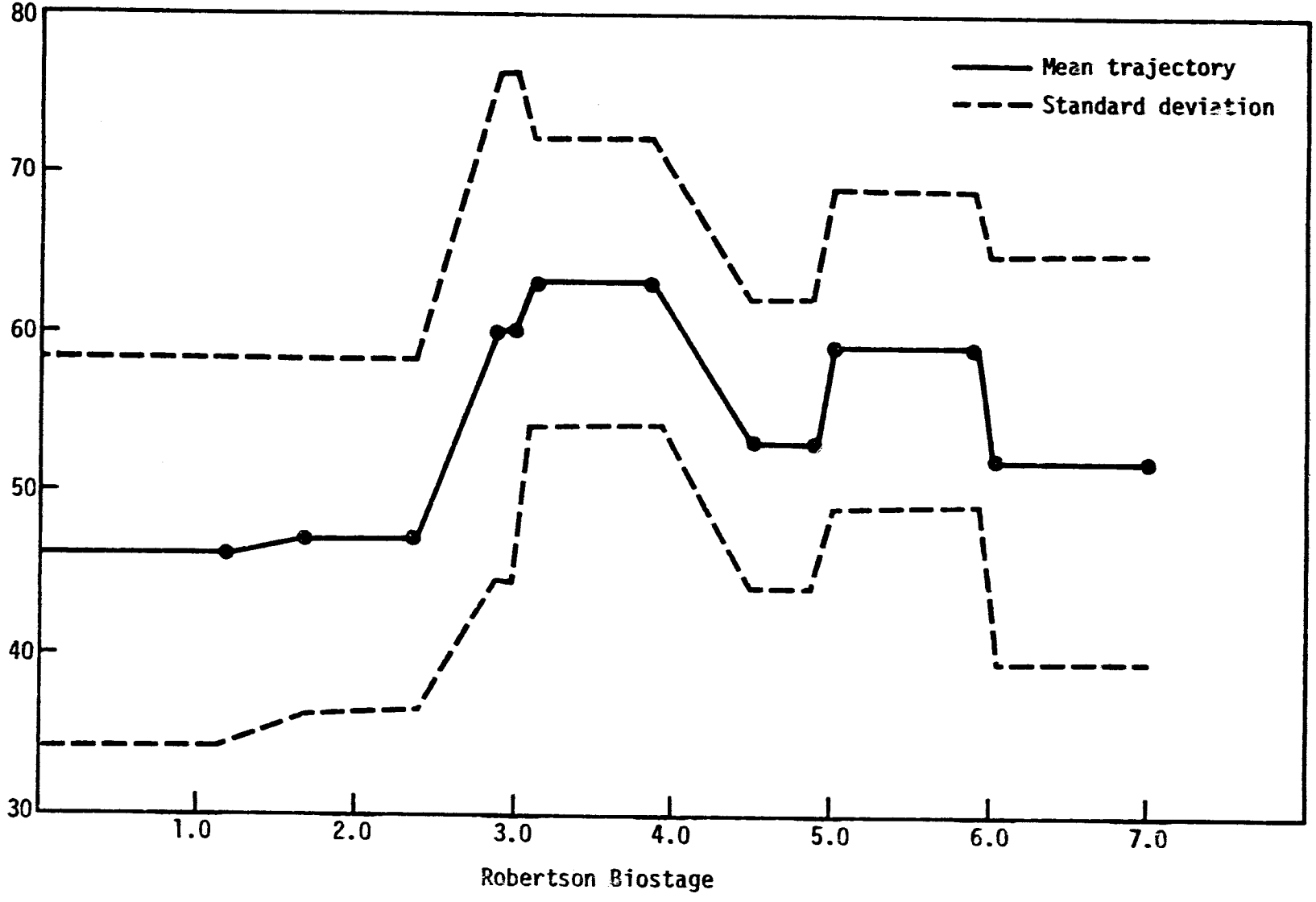


Figure 3-6.- TY brightness key generated from the ground-truth labels.

TABLE 3-4.- ACCURACY OF EXTENSION WITH UPDATED KEYS

(a) Distribution of LIST labels in classification of 24 TY segments with Phase III trained discriminant and updated keys

Ground-truth label	LIST label		
	Small grains	Nonsmall grains	Obvious nonagriculture
Small grains	321	739	14
Nonsmall grains	912	1593	359

Statistics:
- PCC = 57.72%
- Omission rate = 70.11%
- Commission rate = 31.84%
- Bias = +7.45%
- Average PCC across segments = 63%
- Standard deviation of PCC = 18.89%
- PCC, given LIST and AI agree = 84.55%
- PCC of LIST on disagreements = 18.8%

(b) Distribution of LIST labels in classification of 19 TY segments with Phase III weights without updated keys

Ground-truth label	LIST label		
	Small grains	Nonsmall grains	Obvious nonagriculture
Small grains	339	612	12
Nonsmall grains	660	1005	246

Statistics:
- PCC = 55.32%
- Omission rate = 64%
- Commission rate = 34.54%
- Bias = +1.7%
- Average PCC across segments = 57.13%
- Standard deviation of PCC = 20.14%
- PCC, given LIST and AI agree = 81.07%
- PCC of LIST on Disagreements = 18.79%

TABLE 3-5.- RESULTS OBTAINED BY REMOVING BRIGHTNESS KEYS

[Distribution of LIST labels for 24 TY blind sites, classified from Phase III training]

Ground-truth label	LIST label		
	Small grains	Nonsmall grains	Obvious nonagriculture
Small grains	465	595	14
Nonsmall grains	694	1511	359

Statistics:
PCC = 64.18%
Omission rate = 55.4%
Commission rate = 27.07%
Bias = +2.7%
Average PCC across segments = 64.67%
Standard deviation of PCC = 16.97%

TABLE 3-6.- RESULTS USING ONLY GREENNESS/BRIGHTNESS KEYS

Data used in training	Data set classified			
	Phase III		TY	
	Mean PCC	Standard deviation	Mean PCC	Standard deviation
Phase III	83.78	5.19	63.58	17.6
TY	70.26	17.27	82.42	10.26

TABLE 3-7.- RESULTS USING ONLY GREENNESS KEYS

Data used in training	Data set classified			
	Phase III		TY	
	Mean PCC	Standard deviation	Mean PCC	Standard deviation
Phase III	81.89	8.71	65.74	16.87
TY	72.62	20.18	77.24	12.8

TABLE 3-8.- EXTENDABILITY ACHIEVED USING ANALYST KEYS ONLY

(a) Results

Data used to train discriminant	Data classified					
	Phase III			TY		
	Overall PCC	Mean PCC	Standard deviation	Overall PCC	Mean PCC	Standard deviation
Phase III	73.7	73.86	15.69	59	59.15	23.76
TY 68.5	68.55	18.20	74	73.64	21.06	

(b) Probability of agreement of machine classified label and analyst label (classified using only AI keys)

Data used to train discriminant	Data classified	
	Phase III	TY
Phase III	0.567	0.637
TY	.672	.871

TABLE 3-9.- TRAINING AND TEST ACCURACY OF KEYS APPLIED TO THE TY DATA

Data set	Mean PCC	Standard deviation of the PCC
Greenness and brightness keys		
Training data	81.52	10.30
Test data	84.24	10.63
Greenness keys only		
Training data	75.87	12.62
Test data	79.97	13.56
AI keys only		
Training data	72.03	15.69
Test data	76.88	30.2

4. SUMMARY

Sample labeling from satellite MSS data was once performed by means of field delineation and labeling. In order to prevent bias due to subject field selection, the photointerpreter was given a specified set of pixels to label. It was observed then that the AI labeling techniques were highly personalized and yielded results that varied considerably. A questionnaire-discrimination approach to labeling was developed to transform labeling from a personalized art to a transferable technology. Experimental results confirm that accuracy obtained using of this technique can match AI accuracy while yielding less variance; however, its lack of adaptability to crop conditions other than those of the test period suggests that additional development is required for year-to-year extendability.

5. REFERENCES

1. Robertson, G. W.: A Biometeorological Time Scale for a Cereal Crop Involving Day and Night Temperatures and Photoperiods. Int. J. of Biometeology, vol. 12, 1968, pp. 191-223.

2. Kauth, R. J.; and Thomas, G. S.: The TASSELLED CAP-A Graphic Description of the Spectral-Temporal Development of Agricultural Crops as Seen by Landsat. Proc. 10th Annual Symp. on Remote Sensing of environment (Ann Arbor, Michigan), Oct. 1-2, 1975.

3. Pore, M. D.: Label Identification from Statistical Tabulation (LIST). LEC-10825, Rev. A, Nov. 1977.

4. Abotteen, R. A.; and Pore, M. D.: Development of Semiautomated Operational Label Identification from Statistical Tabulation (LIST). LEC-11457, Feb. 1978.

5. Pore, M. D.; and Abotteen, R. A.: Label Identification from Statistical Tabulation (LIST) Test and Evaluation. LEC-12381, Nov. 1978.

6. Dennis, T. B.; and Pore, M. D.: Improvement of Green Number and Brightness Keys Used in Label Identification from Statistical Tabulation (LIST). LEC-12714, Oct. 1978.

7. Pore, M. D.; and Dennis, T. B.: Report on the LIST Operational Test in Kansas; 63-1827-3705-07, Aug. 1979.

8. Dennis, T. B.; and Pore, M. D.: LIST Spectral Keys Study. LEC-13456, Aug. 1979.